International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 7 July 2022, pp: 1098-1103 www.ijaem.net ISSN: 2395-5252

Animal species detection and classification framework based on CNN

Monisha Uday, Dr. S Padmashree
Dept, of ECE, GSSSIETW, Mysuru, Karnataka, India
Professor, Dept, of ECE, GSSSIETW, Mysuru, Karnataka, India

Submitted: 10-07-2022 Revised: 17-07-2022 Accepted: 21-07-2022

ABSTRACT— Detection and classification of animal species is the initial step . Object recognition methods provide recognition of objects of a particular target class in a particular image and assign each object to the corresponding class marker. These practices are proliferating in many ways in network design, system readiness, and capacity rationalization. This paper focuses on animal species recognition as a first step in mitigating the negative effects of encountering wildlife, human and wildlife vehicles in remote areas and highways. To automate the detection process while maintaining robustness to blurring, partial occlusion, lighting, and pose changes, the YOLOv3 object detection framework is used to efficiently detect animals in images.

Keywords—Yolo V3, Wildlife, object detection

I. INTRODUCTION

Identification and Categorization of creature species is a district that wants extraordinary strategies as it reduces the issues of untamed life highway incidents inciting passing and wounds and helps individuals with understanding assortment better. Animals often follow, causing death and wounds in most people. Animal attack replays vary depending on where you live. In the United States, a check of 2,000,000 animals' assaults on individuals are recorded consistently [1] . Curiously, Tanzanian and American analysts' reports provide that Animal attack on individuals increase from 1990 to 2005. It also communicated that somewhere near 563 residents were pursued and eaten during this time. Tracker animal is a casual word portraying a kinds of animal that pursuits individuals as prey. Lion, for example, are known to kill a greater number of individuals than another species of this sort, as moreover recorded by Nowak et al [2] . Lions have not been overlooked, as they have also been reported to attack people living in common places both at night and during the day in search of food. Warrell [3] states that animal attact cause countless passing each year all through the world. Regardless, it doesn't make the feeling that every organization tracks animal related passing. Most animals don't pursue individuals regularly, except for tigers, other explicit animals feed on crippled, neglectful, or dead and decayed human. When animals become used to individuals or are truly starved, they could pursue pets, trained creatures, and human. Animal attacks are by and large typical during the night due to hankering. Subsequently the animals meander about searching for food. Amazing and perfect techniques are supposed to recognize, request, and clasify animals even more beneficially, thusly hindering animal vehicle accidents, following animals. One rapidly creating and normal PC vision task is Object Detection. Significant techniques like CNNs have of late been shown to accomplish different picture understanding the extraordinary execution of continuous investigation outcom. Existing thing identifiers can be in like manner be gathered into two sorts, one-stage and two-stage locaters. Predict the target bobbing box for a single tier locator using different anchor sizes. For example, RetinaNet [4], SSD [5], YOLO90 0 0 [6], YOLOv2 [7] are similar to the early days of faster CNN. [The 8th]. Two-level locators such as Faster R-CNN [9], R-FCN [9], and FPN [10] fully integrate the Common Proposal Association (RPN) to plan the network. The two-level structure is a common point of researcher convergence in today's object identifiers because of its awe-inspiring accuracy. PC Vision (CV) approaches for animal distinguishing proof are crucial, adding to other animal affirmation and classifi-cation approaches in handling exceptional issues like untamed life disasters and risked species [11]. CV has transformed into a by and large used gadget in imagery, prosperity, vehicle arranging, and robots in present day culture [12]. These uses are given up to recognize objects, similar to limitation, area, and Classification [13]. Enormous differentiations in shape and assortment appearance from changed objects of a comparative class are essential issues impacting object identifiers' viability in im-age taking care of [14-17]. The affirmation of a couple of animals potentially suggests that one animal requirements to isolate between others of a comparative form. These problems had accomplished a couple of human issues before development was

used [18] . More-got done, the ID of species heads has clear obstacles, as the substance of animals moves generally from the embodiment of individuals. An exact animal may rapidly take up different designs and assortments to take up an attempting to perceive. The continuous progression of mind networks has on a very basic level upgraded these latest visual affirmation systems.
In this work, we center around further developing execution in exactness and recognition speed of creature species recognizable proof and confinement. This is accomplished by improving the separated highlights through adding deformable YOLO v3 model. Apparently, there is no current work that involves this method in animal recognition in the different weather condition.

II. DATABASE

In our investigation, We used three datasets: (1) the snapshot serengeti dataset [58], (2) the dataset provided by BCMOTI, and (3) the snapshot Wisconsin dataset [59]. Preview Serengeti is a dataset of animal species from Africa (Serengeti National Park, Tanzania). 712,158 images of 7 species (lion, zebra, buffalo, giraffe, fox, deer, elephant) were selected. The BCMOTI dataset contains 53,000 images of eight species commonly found on Canadian roads and remote areas (bears, elks, mousses, deer, cougars, mountain goats, foxes, and wolves). The snapshot Wisconsin dataset was collected in North America using 1037 camera traps installed in the Wisconsin woodlands. It contains 500,000 images of various animal species, and 6 animal species (bear, deer, elk, wolf, fox) were selected. Encounters between these animals also lead to exorbitant breakdowns on the highway. These animals are a portion of the time related with terrible direct encounters with individuals too.

In these three datasets, the classes are imbalanced, which is an issue that needs to be addressed later. Photos are classified by human experts as blanks or animal species names. The resolution of the photos in the dataset ranges from 512 * 384 to 2048 * 1536 pixels. Snapshot Wisconsin,BCMOTI and,Seragati contrast in various points, for instance, dataset size and camera position.

III. METHODOLOGY

The YOLO model was given by Joseph Redmon in his research “You only look once, Unified, Real-time object detection”. The instrument for the calculation utilizes the utilization of a solitary brain network that snaps a picture as an Input and endeavors to anticipate bouncing boxes and class names for each jumping box straightforwardly. Albeit this offered less pre-dictive exactness, which was for the most part because of more limitation blunders, it bragged speeds up to 45 edges each second and up to 155 casings individual on speed improved adaptations of the model [1].

To supply more accentuation a case will be given. For instance, a picture might be partitioned into a 7×7 matrix and every cell in the framework might foresee 2 jumping boxes, coming about in 94 proposed bouncing box expectations. The class probabilities map and the jumping boxes with confidences are then joined into a last arrangement of bouncing boxes and class names.

The YOLO was not without deficiencies, the calculation had various limits be-reason for the quantity of lattices that it could run on as well as a few different issues which will be promotion dressed in this way. Right off the bat, the model purposes a 7 ×7 network and since every matrix can distinguish an item, the model confines the greatest number of articles perceptible to 49. Furthermore, the model experiences what is known as a nearby location model, since every lattice is just fit for identifying one article, in the event that a network cell contains more than one item distinguishing it will not be able. Thirdly, an issue could emerge on the grounds that the area of an item may be in excess of a matrix, in this way, there exists a likelihood that the model could distinguish the article at least a couple of times [2]. Due to the previously mentioned issues experienced while running YOLO, it was genuinely clear that restriction mistake and different issues of the framework should have been tended to. Because of that, YOLOv2 was made as an improvement to manage the issues and questions presented by its ancestor. In this manner, restriction mistakes as well as blunders of genuine were essentially tended to in the new adaptation.

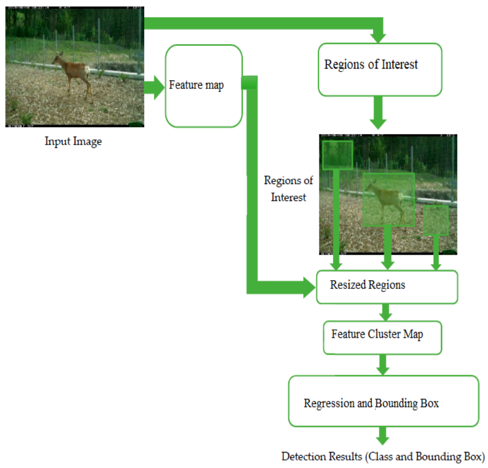

Fig .1.Flow chart for animal detection using YOLO v3

A. Structure:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 7 July 2022, pp: 1098-1103 www.ijaem.net ISSN: 2395-5252

Consequences be damned is carried out as a convolution brain organization and has been assessed on the PASCAL VOC location dataset. It comprises of a sum of 24 convolutional layers followed by 2 completely associated layers. The layers are isolated by their usefulness in the accompanying way:

- Initial 20 convolutional layers followed by a normal pooling layer and a completely associated layer is pre-prepared on the ImageNet 1000-class grouping dataset.
- The pretraining for grouping is performed on dataset with goal 224 ×224.
- Layers include 1x1 drop layers and 3x3 convolution layers.
- The last four convolutional layers followed by two fully connected layers are added to prepare the organization for object discovery.
- Object location requires more granular detail consequently the goal of the dataset is knock to 448 ×448.
- The last layer predicts the class probabilities and bouncing boxes.

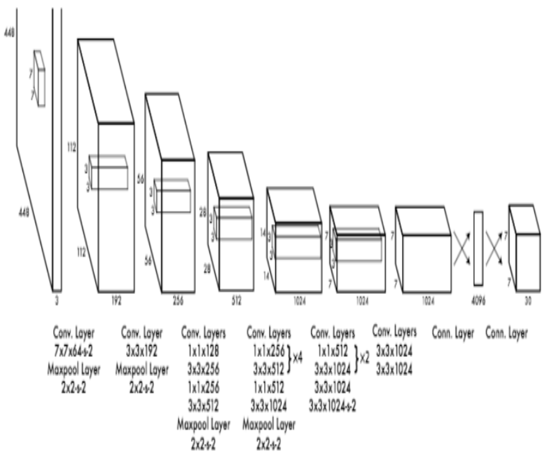

Fig.2. YOLO Structure

B.Working Principle :

Utilizing the given info picture by the client the framework will shows the result with text and exactness. Later the picture is broken into different lengths and widths in light of the given information picture. Here for the acknowledgment of picture, YOLOV3 model is utilizing recognizer profound learning bundle. The recognizer class empowers the model to perceive with the assistance of an internet based apparatus that is google API which is utilized for perceiving various pictures. This model shows the name of creature for the result picture. To show the result picture we utilized the bundle which plot a crate around the creature.

C.Psedo Code:

- Step-1: Start
- Step-2: Reading the picture utilizing OpenCV bundle.
- Step-3: If the model got a picture then the model goes to the perceiving stage. Or, in all likelihood requests that the client give picture as an information.
- Step-4: Darknet comes into the usefulness.
- Step-5: Recognition of picture with name of creature in the wanted language for client.
- Step-6: Displays the result picture.
- Step-7: Repeat the means 1,2,3,4,5,6.
- Step-8: End

Figure 2 stream outline as First we are perusing the pictures from the client then the pictures has been shipped off highlight extractor to isolate the picture into various scales these scales has been associated with the finder gradually ease in this the pictures are been perceived and afterward the jumping boxes will come to usefulness. In this element extractor is utilized for picture handling and for identifies pictures.

First and foremost, the darknet is cloned into store and afterward empowers the GPU and OPENCV, where GPU is utilized for running the model continually and afterward it moves to Compiler Driver which assists with concealing the coordinated work and after that it fabricates the darknet and it downloads the pretrained dataset. Then we will peruse the picture and afterward it transfers the records to darknet and consequently it downloads the documents from darknet then it mounts the google drive and predicts the picture with name and exactness.

Here, Right result implies the picture which are anticipated right kind of creature name were as Wrong result implies the pictures which are anticipated an alternate name instead of the right name of the given information picture. No result implies foreseeing the given information images can't.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 7 July 2022, pp: 1098-1103 www.ijaem.net ISSN: 2395-5252

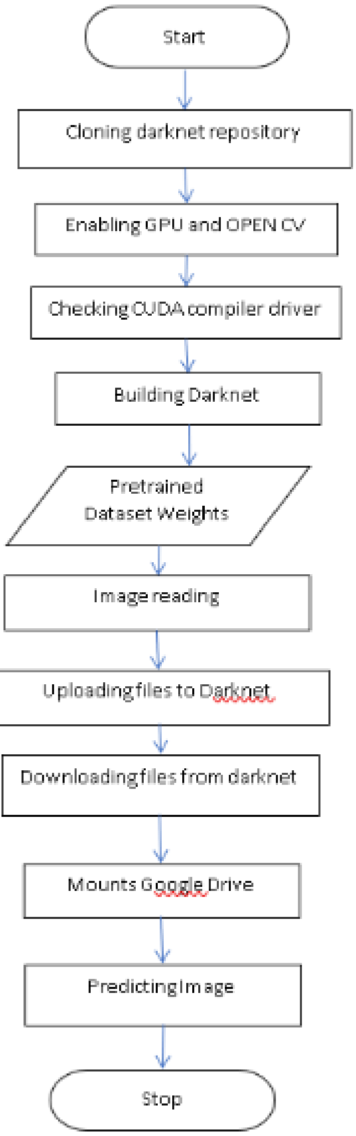

Fig 3. Flow chart

IV. RESULTS AND DISCUSSION

The examination was led on a work station furnished with Windows 10. The proposed system was composed utilizing accessible libraries that suggested framework was developed in Python 3.6 utilising free libraries such as numpy 1.16.5 and scikit-learn 0.21.3. Keras 2.31 together with tensor flow-gpu was used to provide a major neural network framework that was practical for Python.

Moving learning was evaluated using constrained scheduling resources and resources. Move learning was intended to alter a specific model for the typical job while taking into account prior models. Each of the 200 readiness assessments took roughly 17 hours, and the best age for the model was chosen based on testing results following the aggregate setup incident. The Intersection-over-Union (IoU) of sureness and bouncing keep the association models were given a breaking value of 0.5.

Figure 4. the Snapshot Wisconsin dataset with 0.6 percent FNR, Figure 4 demonstrates that by using deformable YOLO V3, exactness and mAP of discovery are both 97.6 percent even in fig 4. The system was developed for the Snapshot Serengeti dataset using a larger preparation set than BCM.

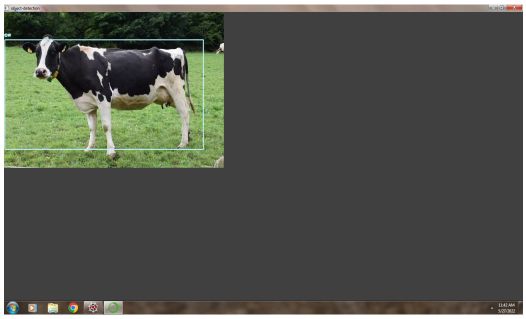

Fig. 4.Detection of cow in normal image

In figure 5 since the majority of the data in this image set were obtained in night mood with low resolution, the exhibition yoylo v3 is 93.3 percent, the tiniest information in this work.

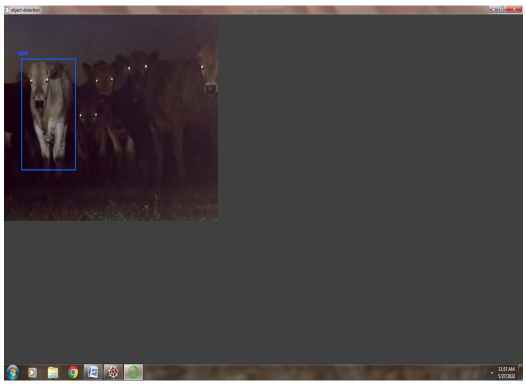

Fig .5. Detection f Animal in poor resolution

Figure 6 demonstrates that, when applied to the Snapshot Wisconsin dataset, YOLO v3 accuracy is 97.6%. The algorithm was trained on a bigger training set before being applied to the Snapshot Serengeti dataset. This demonstrates the value of having a sizable training set with many of examples for each class.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 7 July 2022, pp: 1098-1103 www.ijaem.net ISSN: 2395-5252

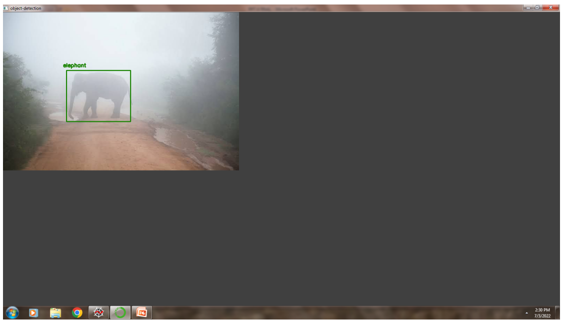

Fig .6. Detection in animals in fogg

Along the confidence score for each class, the picture findings in Figure 7 demonstrate the deformable Mask R-CNN can recognise and seperate both single and numerous animal species. YOLO v3 recognises animal species more quickly and accurately.

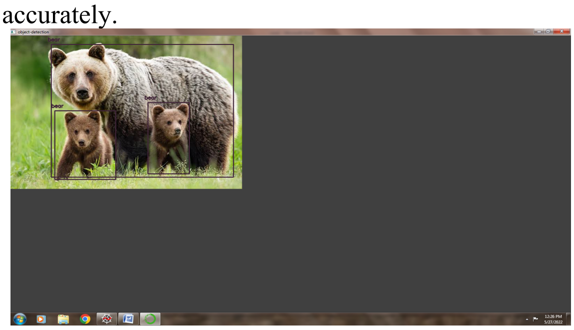
Fig .7. Detection in multiple images

Figure 8 demonstrates that, when applied to the Snapshot Wisconsin dataset, YOLO v3 accuracy is 97.6%. The algorithm was trained on a bigger training set before being applied to the Snapshot Serengeti dataset. Here the detection of multiple animal in single frame is detect.

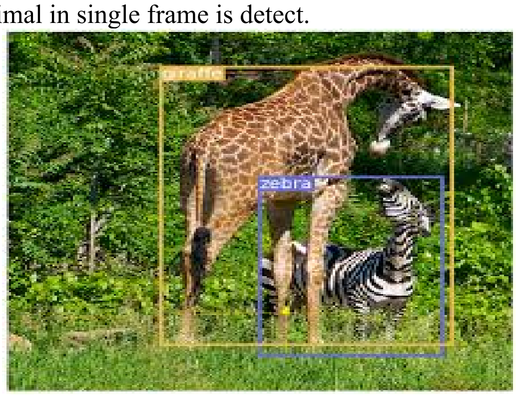

Fig.8. Detection in multiple animals of different breed

V. CONCLUSION

In this paper profound learning-based object identification model utilizing YOLO v3. Then, at that point, these models are assessed on creature pictures from three datasets for high accuracy, constant creature species identification. Then, the exactness and speed execution of creature species identification are given in the wake of improving the separated elements by utilizing

D-CNNs. The outcomes show that YOLO v3 is the ideal decision continuously creature species location and it can accomplish the best exhibition in precision, and mAP. Besides, YOLO v3 gave the promising outcomes to distinguishing creature species at a wide scope of lighting, shadows, and weather pattern.

In future work, we mean to identify more modest creature species which is one of the significant difficulties of creature species discovery and to examine improvement by decreasing FNR. Moreover, intend to plan an effective creature finder by working on the exactness of creature species ID and confinement in sufficiently high speed to be applied continuously applications. To get higher exactness, it is necessary to separate more huge highlights, improve pre-and post-handling techniques, settle the irregularity class issue, oblige lopsidedness constantly pictures, and upgrade characterization certainty.

ACKNOWLEDGMENT

We would like to thank our Principal, Dr. M shivakumar and Head, Dr. Rajendra R Patil for his support for completing the work also we also thank your PG coordinator and project guide , Dr. Shyamala C and Dr. S Padmashree for their continuous support. I also thank DST-CURIE for there support and for providing financial assistance to publish this paper.

REFERENCES

[1] Michal Maj “Object det ect ion and Image Classificat ion with YOLO” KDnugget s , Sept ember 2018.
[2] Ayoosh Kathuria “What’s new in YOLO v3?. A review of the YOLO v3 Object” Towards data science , April 23rd 2018.
[3] Et han Yanjia Li “Dive Really Deep into YOLO v3: A Beginner’s Guide” T owards dat a science, December 31st 2019.
[4] Joseph Redmon , Ali Farhadi “ YOLO v3: An Incremental Improvement” University of Washington, April 8 th 2018.
[5] Manogna Mant ripradaga “ Digging deep into YOLO v3 – A handson guide” T owards dat a science, August 16 th.
[6] Joseph Redmon “Darknet-53 Explained” Cornell University, April 2018, https://paperswithcode.com/method/darknet-53.
[7] Ayoosh Kathuria “ How to implement a YOLO v3 Object det ector from Scrat ch in P yTorch” KDnugget s, May 2018.

[8] S. Sandeep Kumar, Shrut i Mishra, R. Sai Sandeep, U. Sai Ravi Teja, Pradeep Kumar M, M. Shrut i, K. Shravya "Deep learning based image recognition for vehicle number information"Internat ional Journal of Innovat ive Technology and Exploring Engineering, Volume: 08, Issue: 8S2, June 2019.

[9] R. Abinaya, Lakshmana Phaneendra Maguluri, S. Narayana, Magant i Syamala "A Novel Biomet ric Approach For Facial Image Recognition using Deep Learning Techniques" Int ernat ional Journal of Advanced Research in Engineering and Technology (IJARET), Volume: 11, Issue: 9, September 2020.

[10] P. Meghana, S. Sagar Imambi, P. Sivateja, K. Sairam "Image Recognition for Aut omat ic Number P late Surveillance" Internat ional Journal of Innovat ive Technology and Exploring Engineering (IJITEE), Volume: 08, Issue: 04, February 2019.

[11] Pabbisetty Nikhitha, Palla Mohana Sarvani, Kanikacherla Lakshmi Gayat hri, Dhanush P arasa, G Yedukondalu "Det ect ion of Tomat oes Using Art ificial Intelligence Implement ing Haar CascadeTechnique" Int ernat ional Conference on Communicat ion, Comput ing and Electronics Systems, 2020 – Springer.

[12] Sai Manvitha Chittajallu, Navya Lakshmi Deepthi Mandalaneni, Dhanush P arasa, Shahana Bano " Classification of Binary Fracture Using CNN" Global Conference for Advancement in Technology (GCAT), October 2019.

[13] Shahana Bano Vishal P ., Snigdha L.K. "An Efficient Face Recognition Syst em using Local Binary Patt ern" Internat ional Journal of Recent Technology and Engineering (IJRTE), FebruaryG. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[14] Gupta, S.; Chand, D.; Kavati, I. Computer Vision based Animal Collision Avoidance Framework for Autonomous Vehicles. Inf. Process. Manag. Uncertain. Knowl.-Based Syst. 2021, 1378, 237–248.

[15] Gupta, S.; Chand, D.; Kavati, I. Computer Vision based Animal Collision Avoidance Framework for Autonomous Vehicles. Inf. Process. Manag. Uncertain. Knowl.-Based Syst. 2021, 1378, 237–248.

[16] Saxena, A.; Gupta, D.K.; Singh, S. An Animal Detection and Collision Avoidance System Using Deep Learning. Adv. Graph. Commun. Packag. Technol. Mater. 2021, 668, 1069–1084.

[17] Yilmaz, A.; Uzun, G.N.; Gurbuz, M.Z.; Kivrak, O. Detection and Breed Classification of Cattle Using YOLO v4 Algorithm. In Proceedings of the 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Kocaeli, Turkey, 25–27 August 2021; pp. 1–4.

[18] Sato, D.; Zanella, A.J.; Costa, E.X. Computational classification of animals for a highway detection system. Braz. J. Veter-Res. Anim. Sci. 2021, 58, e174951.

[19] Saxena, A.; Gupta, D.K.; Singh, S. An Animal Detection and Collision Avoidance System Using Deep Learning. Adv. Graph. Commun. Packag. Technol. Mater. 2021, 668, 1069–1084.

[20] Labeled Information Library of Alexandria: Biology and Conservation (LILA BC). Available online: http://lila.science/datasets/snapshot-serengeti.[SnapshotSerengeti] (accessed on 27 August 2020).

[21] SnapshotWisconsin, A Volunteer-Based Project forWildlife Monitoring. Available online: https://dnr.wisconsin.gov/topic/research/projects/snapshot.[SnapshotWisconsin] (accessed on 1 May 2020).